research at work

NPRST

Navy Personnel Research, Studies, and Technology
5720 Integrity Drive • Millington, Tennessee 38055-1000 • www.nprst.navy.mil

NPRST-TR-14-2 **February 2014**

Validation of Loss and Continuation Rate Models to Support Navy Community Management

Ping Ying Bellamy, PhD
Tanja F. Blackstone, PhD
Navy Personnel Research, Studies, and Technology

NPRST-TR-14-2
February 2014

Validation of Loss and Continuation Rate Models to Support Navy Community Management

Ping Ying Bellamy, PhD
Tanja F. Blackstone, PhD
Navy Personnel Research, Studies, and Technology

Reviewed by
David M. Cashbaugh

Approved and released by
David M. Cashbaugh
Director

Navy Personnel Research, Studies, and Technology
Navy Personnel Command
5720 Integrity Drive
Millington, TN 38055-1400
www.nprst.navy.mil

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** (From-To)
05/02/2014	Technical Report	Oct 2012 -January 2013

4. TITLE AND SUBTITLE	**5a. CONTRACT NUMBER**
Validation of Loss and Continuation Rate Models to Support Navy Community Management	
	5b. GRANT NUMBER
	5c. PROJECT ELEMENT NUMBER
6. AUTHORS	**5d. PROJECT NUMBER**
Ping Ying Bellamy, PhD Tanja F. Blackstone, PhD	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATIONS NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Navy Personnel Research, Studies, & Technology (NPRST/BUPERS-1) Bureau of Naval Personnel 5720 Integrity Drive Millington, TN 38055-1000	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
Dr. Kip Krebs Office of Naval Research, Code 34 One Liberty Center, 875 N. Randolph Street, Suite 1425 Arlington, VA 22203-1995	
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Continuation rate models are used to forecast the number of individuals available in a given time period by paygrade, years of service, and skill group. Forecasted manpower inventory, the number of individuals available in a given time period, are derived from continuation rates, which in turn inform the advancement and gains modules used within the Department of Defense. As such, the accuracy of continuation rates is critical to these related functions.
The calculation of continuation rates can be based on business rules, statistical analysis, or a combination of both. Using enlisted personnel data from 2003-2011, we separately estimate and compare four different methodological approaches for a subset of enlisted skill groups. We validate each of the methods based on absolute error differences between forecast results and data observations. Of the four methodologies, the findings indicate significant differences in performance across models. Variation in model performance may be affected by sample size and factors not directly captured by the models, but having indirect effects on continuation rates, such as policies governing skill group contracts, promotions, bonuses, and or economic events. Based on forecasted results, we provide a detailed comparison of model performance.

15. SUBJECT TERMS

Continuation rates, forecasting, advancement, gains

16. SECURTIY CLASSIFICATION OF:			**17. LIMITATATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
UNCLASSIFIED			UNLIMITED	56	Wendy Douglas
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED			(901)874-2218

Foreword

The Office of Naval Research's enabling capability program sponsored the development of a prototype computer simulation application to assist community managers and the Community Management Metrics, and Analytics Branch in forecasting future personnel inventories, by simulating Sailor behavior based on input/controls/influencers on accession planning, separation management, advancement/promotion planning, authorizations, and incentive policies.

Integrated Manpower, Personnel, Analysis, Computer Tool for the active-duty component (IMPACT-AC) is a simulation application composed of six modules which represent workforce planning actions, i.e., Enlisted and Officer strategic planning, all Navy Enlisted and Officer reconciliation, and Enlisted and Officer execution monitoring. These modules are based on business rules and use statistical models to simulate, estimate, and forecast the facets entailed in managing Navy workforce needs at the appropriate skills and experience levels. A primary module embedded in IMPACT is the continuation rate module, which is a means of forecasting the individuals available in a given time period by their paygrade, years of service, and enlisted management code (i.e., community). The forecasts of future personnel inventories inform other actions, such as advancements and recruiting. As such, the accuracy of continuation rates is critical to the accuracy of the application's output. The objective of this effort is to independently validate five statistical forecasting methods, based on absolute error differences between forecast results and data observations.

This effort is supported by the Office of Naval Research, Capable Manpower Future Capability, Code 34. The point of contact for this effort is Dr. Tanja Blackstone, Navy Personnel Research, Studies, and Technology, (901) 874-4633

DAVID M. CASHBAUGH
Director

Summary

Continuation rate models are used to forecast the number of individuals available in a given time period by paygrade, years of service, and skill group. Forecasted manpower inventory, the number of individuals available in a given time period, are derived from continuation rates, which in turn informs the advancement and gains modules used within the Department of Defense. As such, the accuracy of continuation rates is critical to these related functions.

The calculation of continuation rates can be based on business rules, statistical analysis, or a combination of both. Using enlisted personnel data from 2001-2007, we separately estimated and compared four different methodological approaches for a subset of enlisted skill groups. We validated each of the methods based on absolute error differences between forecast results and data observations. Of the four methodologies, the findings indicated significant differences in performance across models. Variation in model performance may be affected by sample size and factors not directly captured by the models, but having indirect effects on continuation rates, such as policies governing skill group contracts, promotions, bonuses, and economic events. Based on forecasted results, we provide a detailed comparison of model performance.

Table of Contents

List of Figures

List of Tables

Introduction

IMPACT is an integrated suite of simulation models that are designed to facilitate the planning, executing, and monitoring of enlisted and officer management communities. These models allow for a wide variety of within and out-year execution simulations encompassing the entire breadth of community management from accession planning, conversions and lateral transfers to end-strength planning.

Underlying IMPACT are "modules" that use rules or statistical models to simulate, estimate, and forecast the facets entailed in managing Navy manpower and personnel. A primary module embedded in IMPACT is the continuation rate module. Forecasted manpower inventory, the number of individuals available in a given time period, are derived from continuation rates which in turn informs the advancement and gains modules. As such, the accuracy of continuation rates is critical to the related modules.

The calculation of continuation rates can be based on business rules, statistics, or a combination of both. To ascertain the best approach for forecasting continuation rates, the Center for Naval Analysis (CNA), Pinelis (2011), proposed four different methodological approaches. The objective of this effort is to independently validate each of the four methods and based on forecasted results; provide a detailed comparison of model performance.

In Section II, a discussion of each of the models is provided. More detailed descriptions of model assumptions, and the pros and cons can be found in Pinelis (2011) In Section III and IV basic statistics and results are discussed for each model. A discussion of the forecast results is limited to Aviation Structural Mechanic (Safety Equipment) (AME) and Operation Specialist (OS) enlisted community for the first quarter 2008 community. Continuation rate confidence intervals and their usability within the IMPACT system are addressed in Section V. Section VI outlines the data assumptions.

Model Descriptions

Following CNA technical report *Forecasting Navy Continuation Rates: An Exploratory Analysis*, January 2011, NPRST validated the models using data from the AME and OS enlisted communities. CNA proposed four alternative methodological approaches and models to forecast Continuation Rates (CR). The pros and cons of the proposed approaches are described in the CNA technical report.

Model 1: Moving Average Method- This is the method commonly used by BUPERS-3 analysts. It involves taking average continuation rates over the last 3, 5 and 6 years and using them as predictions for the next year's rates. Following CNA, NPRST estimated moving average models for 3, 5, and 6 years.

(a) $M_{t+1} = [(CR_t + CR_{t-1} + \ldots CR_{t-(N-1)})] / N$
where M = forecasted average CR from 3, 5, or 6 years

N = total number of periods. For 3-years Moving Average, N = 12 quarters.

Model 2: Pseudo-Bayesian – This methodological approach overcomes the problems associated with small sample sizes by incorporating cohort specific information with aggregate cohort data or ALLNAV data. This approach requires three inputs; a count of the current inventory by pay grade and enlisted management community, the current continuation rates by pay grade and enlisted management community, and current ALLNAV continuation rates by pay grade and years of service. From these three inputs, a forecasted point estimate of continuation rates can be calculated. [1]

Following Pinelis (2011), let

N is the total predicted inventory for the community using that community's previous continuation rates.

N^* is the total predicted inventory for the community using ALLNAV previous continuation rates.

y_{ijt+1} are the individual cell predictions based on the Enlisted Management Communities (EMC) continuation rates.

y^*_{ijt+1} are the individual cell predictions based on the ALLNAV continuation rates.

r_{ijt+1} are the individual cell EMC continuation rates.

r^*_{ijt+1} are the individual cell ALLNAV continuation rates.

$\widehat{K}$ is a pseudo-Bayesian weight.

λ_{ij} are the prior probabilities obtained from the all-Navy cell distribution.

m^*_{ijt+1} are the final predicted cell counts, which add up to N and reflect both our original predictions and the all-Navy distributions over cells. Dividing individual $m^*_{ijt+1}$by last period's inventory in the same cell gives the predicted continuation rate r_{ijt+1} .

Using the notation above, then calculate the predicted EMC inventory by multiplying current EMC inventory by current EMC continuation rates and then summing by cell and total.

i. $y_{ijt+1} = y_{ijt} * r_{ijt}$

ii. $\sum_{i,j} y_{ijt+1} = N$

Next find predicted EMC inventory using the ALLNAV continuation rates:

iii. $y^*_{ijt+1} = y_{ijt} * r^*_{ijt}$

iv. $\sum_{i,j} y^*_{ijt+1} = N^*$

Based on the ALLNAV predicted inventory, N*, find the prior cell probabilities:

v. $\lambda_{ij} = y^*_{ijt+1}/N^*$

---

[1] Cohort is defined as a particular enlisted management community. There are approximately 187 enlisted management communities in the Navy. The sum of all communities is referred to as ALLNAV.

From (vi) find the pseudo-Bayesian weight, $\widehat{K}$:

vi. $$\widehat{K} = \frac{N^2 - \sum_{i,j} y_{ijt+1}^2}{\sum_{ij}\left(y_{ijt+1} - N\lambda_{ij}\right)^2}$$

Insert $\widehat{K}$ into (vii) to find the predicted cell counts, m_{ijt+1}^{*}

vii. $$m_{ijt+1}^{*} = \frac{N}{N+\widehat{K}}\left(y_{ijt+1} + \widehat{K}\lambda_{ij}\right)$$

Dividing m^{*}_{ijt+1} by current inventory, y_{ijt}, gives the new continuation rate

viii. $$r_{ijt+1} = \frac{m_{ijt+1}^{*}}{y_{ijt}}$$

Model 3: The Combination Method combines the information from 3, 5, or 6 year moving average continuation rates into the pseudo-Bayesian model. Whereas the pseudo-Bayesian model used information from the immediate previous time period, the combination model incorporates continuation rate information from 3, 5, or 6 years past quarters. The moving average continuation rates replace the pseudo-Bayesian continuation rates, or r_{ijt} , where r_{ijt} is the individual cell EMC continuation rate. Once the substitution has been made, the combination method follows pseudo-Bayesian approach. The most recent ALLNAV continuation rates is used to augment the predictions.

Model 4: Autoregressive method (AR)- This is a statistical forecasting method that uses previous period information in to forecast future periods. Lagged values of CR are used to forecast CR, with forecast accuracy relying on the number of lags chosen.

(b) $CR_t = c + \sum_{i=1}^{p} \beta_i CR_{t-i} + \varepsilon_t$

where c = constant
β - are the estimated coefficients
CR_{t-i}– lagged continuation rates
p – total number of time periods (i.e., lags) considered

CNA estimated the models using 'R', whereas NPRST used SAS. The robustness of the AR methodology is dependent on the number of lag periods, 'p', included in the model. 'R' automatically finds the optimum number of lags, whereas SAS does not have this automated function. While there are statistical tests that can be used to determine the optimal number of lags, these must be applied to each period and statistically tested. As an alternative, the optimal number of lags was determined using Akaike Information Criterion (AIC), the smaller the AIC, the more robust the model.[2] For each PG/YOS combination, separate AR models were estimated, with the number of lags ranging from 0-8.

[2] For an explanation of AIC see http://en.wikipedia.org/wiki/Akaike_information_criterion.

Descriptive Statistic: Differences in Actual and Forecasted Continuation Rates

Provided in Tables 1-6 are basic descriptive statistics for the difference between actual and forecasted continuation rates. Each table provides the total number of observations used for each model type, N. The number of observations reflects each ALLNAV for the observed differences. The mean and standard deviation for the differences are also provided in Tables 1-6.

Using a F-test to compare the variances between the forecasted AME CR values for the AR and 3-year Moving Average (MA) model, the test indicates differences in the variability. Using the information from the F-test, we then use a t-test for unequal variances to determine if there is any statistical difference in the mean CR difference. The results show no statistical difference between the two sample means. Similarly, t-test for differences in the mean between the 3-year combination and pseudo-Bayesian model were conducted. The results show no differential between the actual and forecasted continuation rates across these two models.

Test results are provided in Appendix C. The mean and standard deviation for the difference between actual and forecasted continuation rates for each cohort group is small, indicating no significant difference in the performance between the methodologies discussed.

Table 1
AME 3-Year MA

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	347
MIN	-0.7495
MAX	0.371825
MEAN	-0.01513
STD	0.183364

Table 2
AME Autoregressive

Statistic	r_diff
N	328
MIN	-0.733719604
MAX	0.391670485
MEAN	-0.010283475
STD	0.186645313

AME Pseudo-Bayesian

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	408
MIN	-0.854474965
MAX	0.745428282
MEAN	-0.020080309
STD	0.22226465

AME 3-Year Combination

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	347
MIN	-0.740530576
MAX	0.349907538
MEAN	-0.005345349
STD	0.182896129

Table 3
OS 3-Year MA

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	582
MIN	-0.779757278
MAX	0.495833333
MEAN	-0.005525453
STD	0.179776192

Table 4
OS Autoregressive

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	548
MIN	-1.158443732
MAX	0.757470755
MEAN	0.002906237
STD	0.17259981

Table 5
OS Pseudo-Bayesian

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	627
MIN	-0.891185832
MAX	0.842400158
MEAN	-0.022593145
STD	0.234198304

Table 6
OS 3-Year Combination

Statistic	Difference Between Actual and Forecasted Continuation Rates
N	582
MIN	-0.71580416
MAX	0.484060416
MEAN	-0.000752652
STD	0.178700959

Results

AME/OS Enlisted Community Model Results and Comparisons

Actual first quarter 2008 continuation rates for the AME and OS enlisted management communities are provided in Tables 5 and 13. Following Icosystem (2011) actual CR by PG/YOS/EMC (Enlisted Management Community) were calculated as follows:

(1) $R = 1 - (s/y0)$

where r – is the continuation rate for Sailors by PG/YOS/EMC

s – is the number of separations from this group's beginning inventory during the sampling period. Separations include those individuals who have a change in PG, EMC, YOS or who are a loss to the Navy.

y0 – is the number of individuals in this group (PG/YOS/EMC) at the beginning of the sampling period. Beginning inventory is derived from the previous period ending inventory and accounts for individuals who are a loss to the Navy.

The separation and inventory counts only consider those observations that were on active duty during the previous and current periods. Individuals shown as inactive or in the reserves are excluded in the separation or inventory counts. Inventory counts include all individuals who have a change of PG-YOS status or are a loss to the Navy for that time period under consideration.

The reported continuation rate for a given time period is the rate that individuals by PG/YOS will continue in that following time period. For example, .7207 (see Table 8), is the rate at which individuals in PG E1-E3 with one year of service will continue in 2nd quarter of 2008. Given a 2008 1st quarter beginning inventory of 100, approximately 72 individuals will continue into the second quarter of 2008.

Blank cells,'-', are observed in the actual continuation rate table, indicating there were no observations for the previous quarter. Blank cells for forecasted continuation rates reflect missing observations in that PG/YOS or missing information in at least one previous quarter. As an example, for the 3-year moving average model, a blank cell will be observed if there is missing data in at least one of the twelve previous quarters.

In cases where a '1' was observed, there were no actual or forecasted losses for that PG/YOS cell. Actual continuation rates never exceed '1'; however, in a very small number of cases observed at senior level PGs, pseudo-Bayesian forecasted rates can exceed '1'.[3] It is assumed this is a result of the methodology and/or small sample sizes for these cohorts. For these cells, Icosystems and CNA recommended that continuation rates be set to '1'. The values in Tables 6, 7, 9, and 10 are shown as forecasted by the model as these results provide information on model validity.

The 3-year Moving Average forecasted CRs are displayed in Table 2. The forecasted CR for 1st quarter 2008 was estimated using the average of the previous 12 quarters CR; 1st quarter 2005 - 4th quarter 2007. The 4th quarter 2008 forecasted CRs are estimated using the average of the previous 12 quarters; 4th quarter 2005 - 3rd quarter 2008. Similarly, 5- and 6-year MA forecast estimates were obtained from the previous 20 and 24 quarters respectively for a given forecast period.

[3] See "IMPACT-AC: Pseudo-Bayesian Loss Model Specification" from ICOSYSTEM, dated: 5/4/2011, page.3. If the adjusted forecasted continuation rate, rr, is greater than 1.0, set rr=1.0.

The AR model forecast results are shown in Table 10. As discussed in Section I, the number of previous quarters used to forecast 1st quarter 2008 values vary based on AIC. The lags used for each cohort model are given in Appendix D.

Standard errors for all models are available and can be requested from NPRST. For purposes of illustrating the model performance, the differences between actual and forecasted CRs for the AME 3-year MA and AR models is provided in Table 11, with Figure 1 depicting the distribution of the range differences across all models. From Table 11 the maximum difference is .18. However, the difference for the OS autoregressive and 3-year moving Average models are .4781 and .4852, respectively.

For all the models considered, the distribution of these differences is shown in Figure 1. Overall, there is little difference in performance across these approaches. The interval range for the difference between actual and forecasted continuation rates is largest for the pseudo-Bayesian approach. Based on the width of the range and considering ease of use and implementation, this approach is the least favorable of the models validated.

Table 7
AME Actual Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7
0	.6341				
1	0.6296	.4286			
2	.7973	.6333	.5		
3	.6667	.7907	.7778		
4	.7273	.678	.5517		
5	-	.7429	.7692	1	
6	0.6667	.7273	.7553	.6667	
7		.3333	.7755	.9091	
8		.5	.6222	.8	
9		1	.6667	.625	
10		1	.7826	.5854	
11			.6667	.7436	.5
12			.8889	.8846	1
13			1	.5714	1
14			1	.6667	.5714
15			-	.6667	1
16			.75	.88	.8571
17			1	.6667	.8333
18			1	.7742	.9
19				.7368	.6111
20					.8333

Table 8
AME 3-Year MA Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7
0	.6978				
1	.7207	.7087			
2	.6897	.6193	-		
3	.6274	.5937	.5678		
4	.6435	.6971	.6627		
5	.6529	.6901	.6932	-	
6	-	.6672	.7379	.6496	
7		.6643	.7114	.6395	
8			.7173	.6901	
9			.7227	.7470	
10			.7354	.7333	
11			.7068	.7243	-
12			.8116	.701	.6666
13			.6722	.7331	-
14			-	.7135	.6361
15			-	.7082	.7461
16			.7495	.7200	.7487
17			-	.7420	.7194
18			-	.7468	.7338
19				.6741	.7160
20					.6721

Table 9
AME Autoregressive Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7
0	0.6495				
1	0.711955924	0.630522246			
2	0.786918307	0.60882108			
3	0.604172578	0.625810855	0.536200411		
4	0.709776188	0.645508181	0.700124126		
5	0.522591419	0.647395495	0.860469871		
6		0.691906572	0.75882519	0.60501771	
7		0.622337826	0.734507136		
8			0.745543951		
9			0.736758818	0.740697776	
10			0.688309683	0.711904211	
11			0.686348474	0.643497932	
12			0.70453964	0.701670118	
13			0.67716178	0.575365411	
14			-	0.733771818	0.554842609
15			0.803813293	0.801667698	0.707750496
16			0.542244972	0.720917755	0.702545188
17				0.731277013	0.685264974
18				0.674401708	0.739506162
19				0.70220153	0.696964371
20					0.769426419

Table 10
AME Pseudo-Bayesian Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7
0	0.766441				
1	0.764866	0.745438			
2	0.769141	0.68438	0.543728		
3	0.687118	0.722175	0.744771		
4	0.638209	0.705161	0.685537		
5	0.27237	0.773377	0.726359	0.685421	
6	0.596687	0.773381	0.784617	0.61433	
7		0.791683	0.779421	0.743363	
8		0.630006	0.764536	0.764571	
9		-	0.750856	0.734038	
10		0.594172	0.706979	0.715265	
11			0.671991	0.745323	0.753
12			0.756151	0.766733	0.769145
13			0.742729	0.763895	0.723255
14			0.7996	0.833206	0.825867
15			0.765769	0.853424	0.811989
16			0.72984	0.772427	0.774122
17			0.842931	0.798437	0.811018
18				0.789676	0.75398
19				0.794166	0.798653
20					0.7693

Table 11
AME 3-Year Combination Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7
0	0.692102				
1	0.699267	0.649111			
2	0.706023	0.636871			
3	0.653629	0.666419	0.616787		
4	0.613372	0.653519	0.635544		
5	0.53842	0.68164	0.657404		
6	-	0.701155	0.71031	0.649692	
7		0.673289	0.710298	0.671647	
8			0.717966	0.714588	
9			0.687799	0.689295	
10			0.644329	0.665369	
11			0.646821	0.67046	
12			0.673619	0.682697	0.650092
13			0.690473	0.690441	-
14			-	0.720365	0.706466
15			-	0.783438	0.768336
16			0.65330	0.687754	0.690773
17				0.733221	0.728899
18				0.716997	0.707146
19				0.703595	0.720363
20					0.691099

Table 12
OS Actual Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7	E8	E9
0	.6122						
1	.8287	.6585					
2	.8049	.7327	.4773				
3	.8235	.786	.601				
4	.25	.8545	.759	1			
5	1	.6957	.6951	.5			
6	0.5	.8077	.8191	.8333			
7	-	.7143	.8187	.8125			
8	1	.5	.8333	.7253	.6667		
9	0.5	1	.8743	.7714	.8889		
10		1	.6897	.7054	.4615		
11			.8154	.7315	.6667		
12			.72	.75	.6552		
13			.8519	.8056	.6829	.6667	
14			.88	.7059	.7692	1	
15			.8	.7381	.8039	.9167	
16			.5556	.7538	.7736	1	
17			.9	.6731	.7742	.6667	
18			.75	.8889	.8082	.65	
19			.8	.5849	.806	.7714	
20					.7073	.6765	0
21					.7667	.7143	1
22					.9524	.6667	1
23					.4444	.5385	.8

Table 13
OS 3-Year MA Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7	E8	E9
0	.7222						
1	.6294	.4874					
2	.5573	.6151	.5002				
3	.4880	.5897	.6312				
4	.4483	.5953	.6931				
5	.5147	.6206	.6942	.5911			
6	.5847	.6520	.7320	.6274			
7		.5792	.7268	.6761			
8		.5427	.7107	.6933	-		
9		.5680	.7106	.7021	.6062		
10		.6805	.7283	.7122	.7154		
11			.7047	.7160	.6501		
12			.7228	.7230	.6782		
13			.7142	.7101	.7216	-	
14			.7093	.7236	.7404	-	
15			.7114	.7144	.7081	.7208	
16			.7098	.7238	.7101	.6674	
17			.7229	.7420	.7159	.6468	
18			.7185	.7288	.7330	.6949	
19			.6116	.6866	.7220	.7294	
20					.7179	.7104	.6611
21					.6924	.7075	.7456
22					.7215	.7089	.6687
23					.6771	.6942	.7410

Table 14
OS Autoregressive Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7	E8	E9
0	0.6889						
1	0.7690	0.5192					
2	0.7684	0.7015	0.5492				
3	0.6776	0.6990	0.6753				
4	0.5246	0.7199	0.7183				
5	0.5219	0.6376	0.8516	0.6278			
6		0.5988	0.8370	0.6255			
7		0.6247	0.8658	0.7128			
8		0.5398	0.8461	0.7380			
9		0.5321	0.7081	0.6699	0.6319		
10			0.7616	0.7174	-		
11			0.7974	0.7151	0.7884		
12			0.7485	0.7201	0.6663		
13			0.7023	0.7251	0.6769		
14			0.6903	-	0.7031		
15			0.5942	0.6977	0.7531	0.6667	
16			0.6548	0.7415	0.6863	0.7230	
17			0.7054	0.7911	0.7968	0.6863	
18			0.6277	0.7017	0.8346	0.7011	
19			0.9284	0.7407	0.7999	0.6732	
20					0.7840	0.6251	1.1584
21					0.7567	0.7170	0.7106
22					0.6891	0.7250	0.6741
23					0.6979	0.7198	0.8694

Table 15
OS Pseudo-Bayesian Forecasted Continuation Rates
First Quarter 2008

YOS/PG	E1-E3	E4	E5	E6	E7	E8	E9
0	0.707026						
1	0.794252	0.666501	0.615431				
2	0.79217	0.697672	0.599859				
3	0.749297	0.763392	0.703824				
4	0.711534	0.745754	0.725121	0.786553			
5	0.832635	0.736492	0.688044	0.312855			
6	0.644316	0.773695	0.777804	0.611313			
7	0.107822	0.719106	0.779538	0.759765			
8	0.75861	0.563378	0.790837	0.806995	0.685701		
9	0.896437	0.480584	0.746748	0.767609	0.766939		
10		0.417663	0.742123	0.70704	0.776201		
11			0.699666	0.7265	0.747479		
12			0.805561	0.743629	0.697458		
13			0.753539	0.770159	0.693976	0.805104	
14			0.801257	0.813152	0.798213	0.374223	
15			0.881281	0.866967	0.865202	0.831001	
16			0.755241	0.742316	0.772373	0.694213	
17			0.721056	0.814493	0.794019	0.755305	
18			0.684914	0.833837	0.797992	0.79847	
19			0.630885	0.779031	0.799593	0.751768	
20					0.757871	0.76396	0.871801
21					0.730036	0.779019	0.881958
22					0.848785	0.752021	0.512502
23					0.751927	0.678867	0.867001

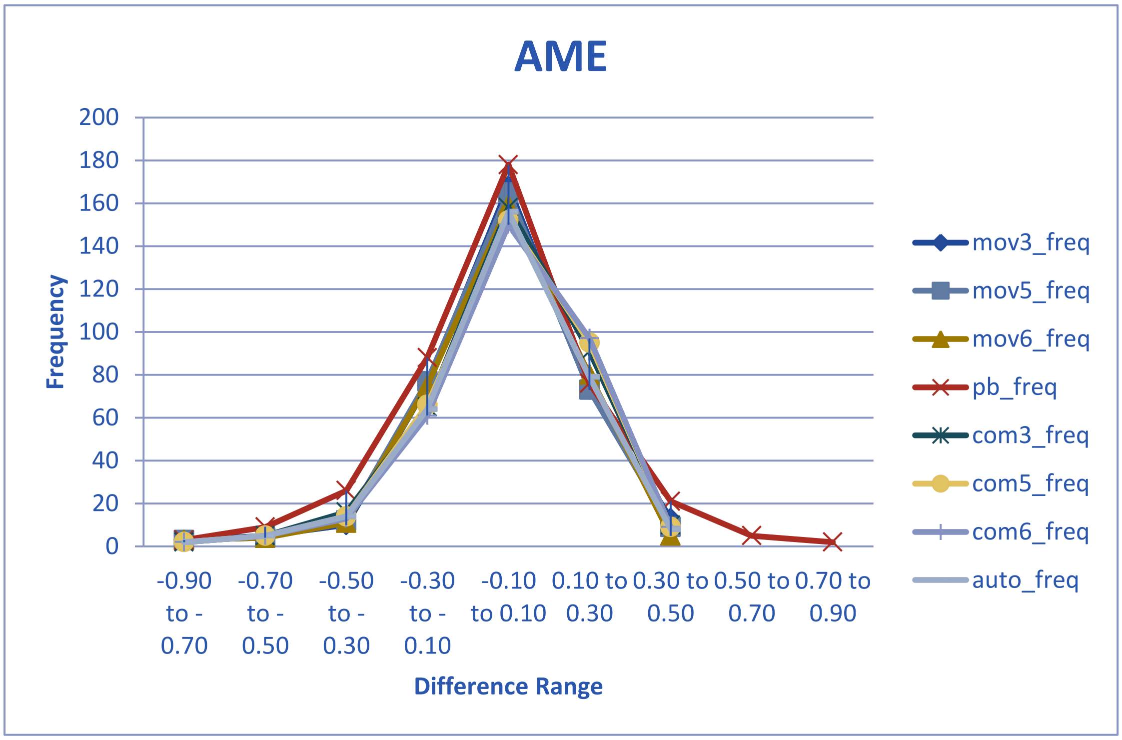
Table 16
OS 3-Year Combination Forecasted Continuation Rates
First Quarter 2008

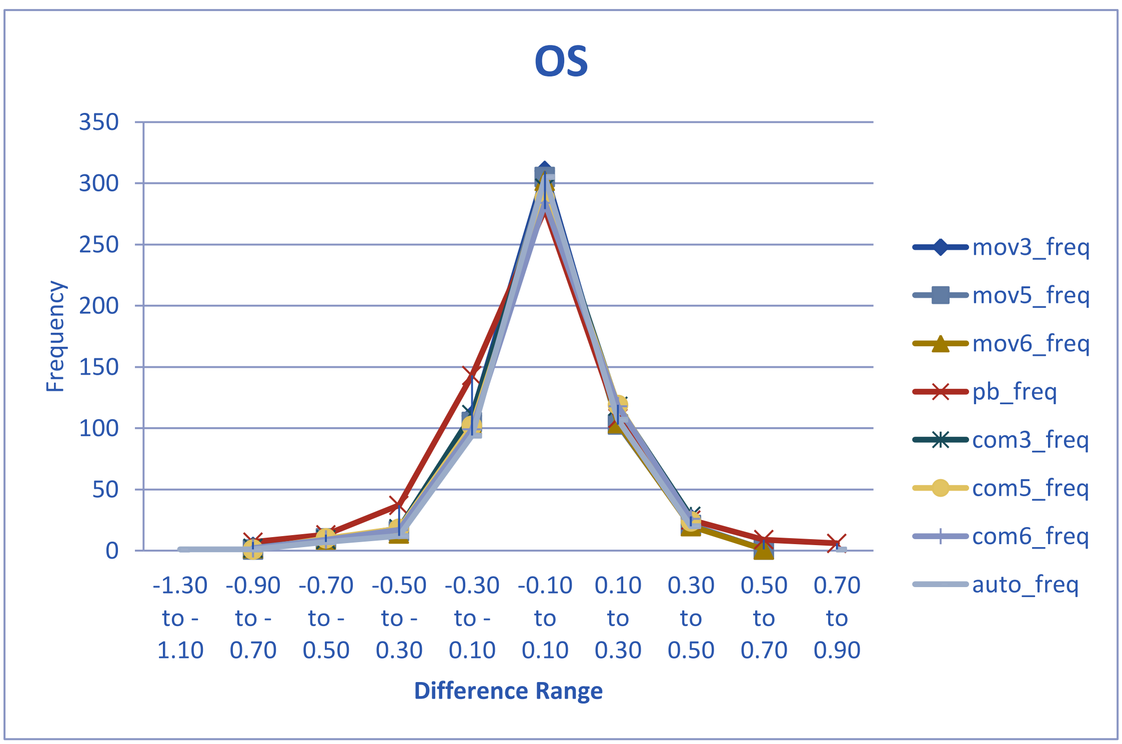
YOS/PG	E1-E3	E4	E5	E6	E7	E8	E9
0	0.703377						
1	0.661362	0.565182					
2	0.631535	0.624289	0.556117				
3	0.571649	0.629579	0.623279				
4	0.529325	0.620452	0.659406				
5	0.560465	0.648676	0.670734	0.569594			
6	0.57144	0.675773	0.716419	0.63601			
7		0.6251	0.715103	0.672295			
8		0.551603	0.711318	0.702211			
9		0.562685	0.694252	0.689845	0.600656		
10		0.560272	0.678066	0.682084	0.655812		
11			0.669304	0.68715	0.659603		
12			0.688255	0.698364	0.660104		
13			0.699754	0.694982	0.682715		
14			0.692729	0.719065	0.722941		
15			0.733847	0.749913	0.736727	0.73259	
16			0.673469	0.700706	0.694494	0.641924	
17			0.696546	0.733827	0.71982	0.675163	
18			0.722185	0.718193	0.715329	0.686747	
19			0.646071	0.693587	0.71815	0.712013	
20					0.701898	0.693036	0.666826
21					0.680166	0.684405	0.71698
22					0.704769	0.692258	0.657043
23					0.676714	0.685299	0.701235

Table 17
AME Autoregressive Model 2008 1st Quarter
Difference between Actual and Forecast Continuation Rates
Select PG/YOS

PG	YOS	Actual	Forecast	Difference
E1-E3	0	0.6341	0.649511	-0.01536
E1-E3	1	0.6296	0.711956	-0.08233
E1-E3	2	0.7973	0.786918	0.010379
E4	4	0.678	0.645508	0.032458
E4	5	0.7429	0.647395	0.095462
E4	6	0.7273	0.691907	0.035366
E5	5	0.7692	0.86047	-0.09124
E5	6	0.7553	0.758825	-0.00351
E5	7	0.7755	0.734507	0.041003
E5	8	0.6222	0.745544	-0.12332
E5	9	0.6667	0.736759	-0.07009
E5	10	0.7826	0.68831	0.094299
E6	6	0.6667	0.605018	0.061649
E6	9	0.625	0.740698	-0.1157
E6	10	0.5854	0.711904	-0.12654
E6	11	0.7436	0.643498	0.100092
E6	12	0.8846	0.70167	0.182945
E6	13	0.5714	0.575365	-0.00394
E6	14	0.6667	0.733772	-0.06711
E6	15	0.6667	0.801668	-0.135
E7	16	0.8571	0.702545	0.154598
E7	17	0.8333	0.685265	0.148068
E7	18	0.9	0.739506	0.160494
E7	19	0.6111	0.696964	-0.08585
E7	20	0.8333	0.769426	0.063907

Figure 1.
Difference between Actual CR and Forecasted CR





Confidence Intervals

The confidence interval is an estimated range of values that will most likely include the unknown population parameter or continuation rate. Tables 11 and 12 provide the upper and lower bounds for the 95% and 99% confidence intervals for the MA and AR AME model. Columns are added to Table 11 labeled ‘within 95%/99% CI’ to show the cases where the confidence interval captures the continuation rate designated by a ‘1’, '0’ otherwise.

The moving average model does not automatically provide a confidence interval. The CI’s provided below, were calculated using the standard CI formula:

(a) Estimated continuation rate +/- (1.96)*(standard error for the MA model)

1.96 is for 95% CI and 2.58 for 99% CI

In Pinelis (2011), CNA proposed estimating the confidence interval for the MA models using:

(b) Estimated continuation rate +/- 1/square root of N

Where N = the size of the community

In most cases the interval estimate from (b) did not capture the population parameter. For N large, the interval estimate is small. As noted in Pinelis (2011), further experimentation is required to determine the best confidence interval.

The non-parametric nature of the combination and pseudo-Bayesian models discussed in Section I does not allow for the calculation of standard confidence intervals. Following Pinelis (2011) an estimated confidence interval was constructed for the combination and pseudo-Bayesian point estimates. Using (b) above, the upper and lower bounds of the confidence intervals are provided in Table 20 for the AME pseudo-Bayesian point estimates. Less than 20% of the confidence intervals capture the point estimate.

The range of values captured by each confidence interval is fairly broad. In some cases, the bounds of the CI are negative or exceed ‘1’. This is a result of (a) and these values can be ignored for purposes of ‘if/then’ analysis. For purposes of use in IMPACT, we recommend choosing starting values within one standard deviation of the forecasted mean continuation rate, compare forecasted value with actual continuation rates over successive time periods, and modify the value drawn from the CI range accordingly. Alternatively, Pinelis (2011), recommends four options to incorporate the interval information. Any option must be considered in view of actual continuation rates.

Table 18
AME 95% and 99% Confidence Intervals
3 Year Moving Average
First Quarter 2008

PG	YOS	95% CI Lower Bound	95% CI Upper Bound	99% CI Lower Bound	99% CI Upper Bound	Within 95% CI	Within 99% CI
E1-E3	0	0.49367451	0.902107	0.429075548	0.966705619	1	1
E1-E3	1	0.467099729	0.974328	0.386874896	1.054552537	1	1
E1-E3	2	0.380509304	0.999031	0.282681913	1.09685827	1	1
E1-E3	3	0.287115864	0.967803	0.179456183	1.075462562	1	1
E1-E3	4	0.262799767	1.024299	0.142358508	1.144740602	1	1
E1-E3	5	-0.02272317	1.328676	-0.23646481	1.542417186		
E1-E3	6					0	0
E4	1	0.279119195	1.138407	0.143211389	1.274315066	1	1
E4	2	0.28945859	0.949312	0.185093965	1.053676973	1	1
E4	3	0.391201706	0.796236	0.327140119	0.860297842	1	1
E4	4	0.415714679	0.978645	0.326679785	1.067679865	1	1
E4	5	0.284146873	1.096157	0.155716748	1.22458682	1	1
E4	6	0.318723386	1.015813	0.20846943	1.126066867	1	1
E4	7	0.26194594	1.066758	0.134654274	1.19404943	1	1
E5	3	0.161254465	0.974444	0.032637766	1.103060612	1	1
E5	4	0.452529155	0.872938	0.386035908	0.939431318	1	1
E5	5	0.475916252	0.910514	0.407178842	0.979251473	1	1
E5	6	0.61840754	0.857484	0.580594381	0.895297443	1	1
E5	7	0.611633109	0.811227	0.580064728	0.842795124	1	1
E5	8	0.607840968	0.826955	0.573185197	0.861610652	1	1
E5	9	0.56416498	0.881336	0.514000238	0.931500348	1	1
E5	10	0.535867543	0.934932	0.472750202	0.998049367	1	1
E5	11	0.325001794	1.088653	0.204220214	1.20943465	1	1
E5	12	0.422353983	1.20086	0.299222881	1.323991405	1	1
E5	13	0.076138618	1.268306	-0.1124184	1.456862885	1	1
E6	6	0.158342843	1.140864	0.002944166	1.296262183	1	1
E6	7	0.342730811	0.936394	0.248835139	1.03028944	1	1
E6	8	0.476816293	0.903458	0.409337204	0.970937362	1	1
E6	9	0.503035958	0.990984	0.425860578	1.068158903	1	1
E6	10	0.503522406	0.963122	0.43083059	1.035814086	1	1
E6	11	0.460218663	0.988555	0.376655301	1.07211812	1	1
E6	12	0.425918217	0.976082	0.3389025	1.063097826	1	1
E6	13	0.400113227	1.066184	0.294765265	1.171532175	1	1
E6	14	0.346543391	1.080614	0.230440317	1.196717513	1	1
E6	15	0.393159784	1.023312	0.293492906	1.122978534	1	1

PG	YOS	95% CI Lower Bound	95% CI Upper Bound	99% CI Lower Bound	99% CI Upper Bound	Within 95% CI	Within 99% CI
E6	16	0.481216623	0.958833	0.405675312	1.034373966	1	1
E6	17	0.519884647	0.96427	0.449599146	1.034555898	1	1
E6	18	0.543172588	0.950505	0.478747595	1.014929794	1	1
E6	19	0.43227282	0.916024	0.355761193	0.992535377	1	1
E7	11					0	0
E7	12	-0.29837062	1.631704	-0.60363752	1.936970854	1	1
E7	13					0	0
E7	14	-0.09921328	1.371436	-0.33181590	1.604038121	1	1
E7	15	0.379340814	1.113053	0.263294441	1.229099739	1	1
E7	16	0.413033099	1.084442	0.30684093	1.190633817	1	1
E7	17	0.416979072	1.021912	0.321300885	1.117590317	1	1
E7	18	0.554464878	0.913223	0.497722597	0.969964808	1	1
E7	19	0.460190957	0.971873	0.37926165	1.052802337	1	1
E7	20	0.210116335	1.13413	0.063971365	1.280274667	1	1

Table 19
AME 95% and 99% Confidence Intervals
Autoregressive Model

PG	YOS	95% CI Lower Bound	95% CI Upper Bound	99% CI Lower Bound	99% CI Upper Bound
E1-E3	0	0.403503	0.895519	0.326201	0.97282
E1-E3	1	0.514549	0.909363	0.452519	0.971393
E1-E3	2	0.525295	1.048542	0.443087	1.13075
E1-E3	3	0.245819	0.962526	0.133216	1.075129
E1-E3	4	0.338886	1.080667	0.222343	1.197209
E1-E3	5	-0.16745	1.212632	-0.38428	1.429458
E4	1	0.311568	0.949477	0.211345	1.0497
E4	2	0.385674	0.831968	0.315557	0.902086
E4	3	0.470537	0.781085	0.421746	0.829875
E4	4	0.383995	0.907021	0.301822	0.989195
E4	5	0.305934	0.988857	0.198639	1.096152
E4	6	0.374338	1.009475	0.274551	1.109262
E4	7	0.179066	1.06561	0.03978	1.204896
E5	3	0.291027	0.781374	0.213988	0.858413
E5	4	0.5219	0.878348	0.465898	0.93435
E5	5	0.687579	1.03336	0.633253	1.087687
E5	6	0.627925	0.889726	0.586793	0.930857
E5	7	0.608595	0.860419	0.56903	0.899984
E5	8	0.583272	0.907816	0.532283	0.958805
E5	9	0.524409	0.949109	0.457683	1.015834
E5	10	0.424387	0.952233	0.341456	1.035163
E5	11	0.363114	1.009583	0.261547	1.11115
E5	12	0.254312	1.154768	0.11284	1.29624
E5	13	0.227417	1.126906	0.086098	1.268226
E5	15	0.411887	1.19574	0.288735	1.318892
E5	16	0.129843	0.954647	0.000257	1.084233
E6	6	0.062347	1.147688	-0.10817	1.318208
E6	9	0.425236	1.056159	0.326111	1.155284
E6	10	0.43731	0.986499	0.351026	1.072783
E6	11	0.463322	0.823674	0.406706	0.88029
E6	12	0.503802	0.899538	0.441628	0.961713
E6	13	0.37697	0.773761	0.314629	0.836102
E6	14	0.543987	0.923557	0.484352	0.983191
E6	15	0.593409	1.009926	0.52797	1.075366
E6	16	0.506846	0.934989	0.43958	1.002255
E6	17	0.56821	0.894344	0.516971	0.945583
E6	18	0.521743	0.827061	0.473774	0.87503

PG	YOS	95% CI Lower Bound	95% CI Upper Bound	99% CI Lower Bound	99% CI Upper Bound
E6	19	0.468392	0.936011	0.394923	1.00948
E7	14	0.013663	1.096022	-0.15639	1.266073
E7	15	0.296549	1.118952	0.16734	1.248161
E7	16	0.40455	1.00054	0.310913	1.094177
E7	17	0.403674	0.966856	0.315192	1.055338
E7	18	0.47383	1.005182	0.390349	1.088663
E7	19	0.46342	0.930509	0.390035	1.003893
E7	20	0.446105	1.092748	0.34451	1.194343

Table 20
AME Margin of Error Confidence Intervals
Pseudo-Bayesian

PG	YOS	Lower CI Bound	Upper CI Bound	Point Estimate Captured in Estimated CI
E1-E3	0	0.734762	0.79812	0
E1-E3	1	0.733186	0.796545	0
E1-E3	2	0.737462	0.80082	1
E1-E3	3	0.655438	0.718797	1
E1-E3	4	0.60653	0.669888	0
E1-E3	5	0.236672	0.308069	
E1-E3	6	0.565008	0.628366	0
E4	1	0.713758	0.777117	0
E4	2	0.652701	0.716059	0
E4	3	0.690496	0.753854	0
E4	4	0.673482	0.73684	1
E4	5	0.741698	0.805056	1
E4	6	0.741702	0.80506	0
E4	7	0.760004	0.823362	0
E4	8	0.598327	0.661685	0
E4	9	-	-	0
E4	10	0.562493	0.625851	0
E5	2	0.512049	0.575408	0
E5	3	0.713092	0.77645	0
E5	4	0.653858	0.717217	0
E5	5	0.69468	0.758038	0
E5	6	0.752938	0.816296	1
E5	7	0.747742	0.8111	1
E5	8	0.732857	0.796215	0
E5	9	0.719177	0.782536	0
E5	10	0.675299	0.738658	0
E5	11	0.640312	0.70367	1
E5	12	0.724472	0.78783	0
E5	13	0.71105	0.774408	0
E5	15	0.767921	0.831279	0
E5	16	0.73409	0.797448	1
E5	17	0.698161	0.761519	0
E5	18	0.811252	0.87461	0
E6	5	0.653742	0.7171	0
E6	6	0.582651	0.646009	0
E6	7	0.711684	0.775042	0

PG	YOS	Lower CI Bound	Upper CI Bound	Point Estimate Captured in Estimated CI
E6	8	0.732891	0.79625	0
E6	9	0.702359	0.765718	0
E6	10	0.683586	0.746944	0
E6	11	0.713643	0.777002	1
E6	12	0.735053	0.798412	0
E6	13	0.732216	0.795574	0
E6	14	0.801527	0.864886	0
E6	15	0.821745	0.885104	0
E6	16	0.740748	0.804106	0
E6	17	0.766758	0.830116	0
E6	18	0.757996	0.821355	1
E6	19	0.762487	0.825845	0
E7	11	0.721321	0.784679	0
E7	12	0.737466	0.800824	0
E7	13	0.691576	0.754934	0
E7	14	0.794188	0.857546	0
E7	15	0.78031	0.843669	0
E7	16	0.742443	0.805801	0
E7	17	0.779339	0.842697	1
E7	18	0.722301	0.785659	0
E7	19	0.766974	0.830332	0
E7	20	0.737621	0.800979	0

Data Assumptions

1. SQL codes provided by Icosystems Corporation were used to populate AME, OS, and ALLNAV data sets The SQL coding contained the rules for gains, losses, paygrades, EMC, inventory counts, etc. The source data for the Icosystems SQL codes is Navy Personnel Data Base (NPDB). Due to the format of historical snapshots and updating rules in the NPDB, extraction of the data from the NPDB proved to be cumbersome. As a result, NPRST opted to use the EMF as the source data.
2. Data contain observations from 1/1999-9/2009, with observations captured quarterly: October 1, January 1, April 1, and July 1.
    - Data for the period from 2001-2007 was used to forecast continuation rates (CR) for 2008 and 2009.
3. Additional variables were added to the AME, OS, and ALLNAV data outputs to account for changes in cohort for inventory. Data variables and definitions are provided in Appendix A.

4. The SQL code calculates years of service, (YOS), as the difference between active duty start date, (ADSD), and date of the observation in the data. Individuals increment into the next YOS at the start of each fiscal year, (FY). Observations, therefore, can be captured in the incorrect YOS. It is possible for an individual to be counted in an earlier/later YOS or the same individual to be observed twice in the same YOS. To mitigate incorrect inventory counts, NPRST used SAS ACT/ACT. SAS ACT/ACT uses ADSD and date of the observation to calculate YOS and increments YOS based on the ADSD and not the start of the FY.
5. Observations were only captured if the SC_IND was coded XFXXX. SC_IND is a variable in the EMF that denotes if an individual is considered active duty, reserves, or inactive. The data used for model validation only captures individuals that are considered active duty; SC_IND = XFXXX.
    - Other sc-ind codes account for individuals still considered in the Navy but not active duty. It is possible for an individual to be active in 1 period, not considered active in period 2 and reappear in the data in period 3. There are very few of these observations in the AME/OS data set, but should be considered if models are expanded to other EMC's or for the overall IMPACT system.
6. In addition, NPRST opted to provide model results based on quarters. CNA report date January 2011 output tables showed PG/YOS, where YOS changes at the start of the FY. The relative duration between YOS could lead to inefficiency in the predicted continuation rates. We have opted to compute CR by quarters.
7. CNA used 'R' in estimating the model. NPRST used SAS. It can be the case that differences in the results are a function of the optimization embedded in the software. The proposed models are 'relatively' standard; differences in results due to estimation software used are unlikely, but possible.
8. Following CNA, continuation rates, (CR) for paygrades, (PG), E1-E3 were combined into one category.
9. For AME, OS, and ALLNAV data we observed YOS for a given PG that exceeded guidelines, (NAVADMIN 056/05). Observations with YOS for a given PG that exceeded guidelines were deleted. See Appendix B for guidelines.
10. The AME data only shows observations for PG E1-E7. There are no E8/E9 observations in the data. Recall that the AME data was created using the SQL queries. (This omission of PG should be reviewed for purposes of the overall IMPACT model.)
    - Per discussion with BUPERS, at the E-8 level, AM and AME personnel may be advanced to AMCS (Senior Chief Aviation Structural Mechanic). At the E-9 level, AMCS and ADCS (Senior Chief Aviation Machinist's Mate) personnel look towards advancement to AFCM (Master Chief Aviation Maintenanceman).
    - Given the estimated time required to verify and recreate the AME data sets, NPRST opted not to modify the existing AME data.
    - The OS data does capture PG's E1-E9.
11. CNA/Icosystems proposed two methods for estimating losses. The first method, (reference Icosystems document, dated May 4, 2011*, IMPACT-AC: Pseudo-Bayesian Loss Model Specification*), is a rules based approached. The second method, (reference Icosystems document, dated August 2, 2012*, IMPACT-AC: Pseudo-Bayesian Loss Model Specification*), uses a regression based approached. For

purposes of NPRST model validation, the rules based approached to calculate losses was used.

Conclusions

The objective of this study was to determine which of four proposed models provided the best forecast accuracy. The findings in this analysis indicate marginal differences in performance across models. The pseudo-Bayesian model gives a wider range of forecasted continuation rates. Whereas the difference range for the AR and MA models is narrower.

In contrast to the preliminary findings discussed in Pinelis (2011), the pseudo-Bayesian forecast intervals, using the margin of error confidence interval calculation, do not capture the true continuation rate most frequently, when compared to the other models. This study shows that the MA and AR models perform better in terms of capturing the true continuation rate. However, as noted in Pinelis (2011) the margin of error confidence interval is a proxy and alternative measures need to be explored.

The findings reported herein should be considered in light of the enlisted management communities, (EMC), chosen for the purpose of validating the models. Variation in sample size across EMC, particularly for EMC's with a small number of observations may lend itself to different results. Model performance may also be affected by factors not directly captured by the models, but having indirect effects on continuation rates, such as policies governing EMC contracts, promotions, bonuses, and or economic events. The key benefit of the pseudo-Bayesian approach is that the method accounts for small sample sizes and by incorporating all Navy information improves the reliability of forecast estimates.

In terms of recommending which of the proposed models is 'best', forecast accuracy, ease of use, costs of implementation, and model maintenance should be considered. To this extent the MA models meets these criteria. However, the MA methodology may not perform well in cases stipulated above. Additionally, it is recommended that the forecast be updated at a minimum of every twelve months.

The loss calculation of the continuation rate was rule based. Improvements in forecast accuracy may be found by incorporating into the models military and demographic variables. Future research, as discussed in Pinelis (2011) includes estimating losses by loss type; attrition, quits, or retirements.

References

Pinelis, Y. K. (2011), *Forecasting Navy Continuation Rates: An Exploratory Analysis*, Center for Naval Analysis, CRM D0023555.A2

Schmitz, E., Kraus, A., Sutton, W. & Pinelis, Y. K. (2012), *The Design of the Execution Monitoring Models for Enlisted and Officer Community Management*, Center for Naval Analysis, DRM-2012-U-001072-SR1.

Icosystem Inc. (2011), *IMPACT-AC: Pseudo-Bayesian Loss Model Specification*, non-published report.

Icosystem, (2012), *IMPACT-AC: Pseudo-Bayesian Loss Model Specification*, non-published report.

Rodney, D., (1991), *A Pseudo-Bayesian Technique for Estimating Continuance in PERMIT, a Personnel Management Integration Tool,* CNA Research Memorandum 90-208.

Bishop, Y., Fienberg, S. & Holland, P. *Discrete Multivariate Analysis: Theory and Practice*. New York: Springer, 2007

Appendix A

Description of Data Elements

Data Element	Description
Time/Date	Date of observation in MM/DD/YY format
SSN	Social Security Identifier
Activedate	Active Duty Start Date (ADSD) – source EMF
Days	Difference between ADSD and date of observation in days
Paygrade (PG)	Enlisted Personnel Rank E1-E9
EMC	Enlisted Management Community
NoChange	1 = if no PG, EMC change from the previous to current period, 0 = if from the previous period to the current period PG or EMC or individual is a Loss to the Navy or a gain to the EMC
Gain	1 = if gain to EMC, 0 = otherwise
Loss	1= if loss to the Navy or EMC, 0 = otherwise
PG_change	1 = if PG changed, 0= otherwise
EMC_Change	1 = if EMC change into community, 0 = otherwise
PG_EMC_Change	1 = if PG and EMC change into community, 0 = otherwise
YOS	Years of Service: difference between ADSD and observation date
YOS_Count_all	Inventory count for all individuals in a given YOS regardless of sc_ind code
PG_count_all	Inventory count for all individuals in a given PG regardless of sc_ind code
EMC_count_all	Inventory count for all individuals in a given EMC regardless of sc_ind code
PG_count_XFXXX	Inventory count for all individuals in a given PG with sc_ind = XFXXX (active duty)
EMC_count_XFXXX	Inventory count for all individuals in a given EMC with sc_ind = XFXXX (active duty)
YOS_PG_count_all	Inventory count for all individuals in a given PG, EMC regardless of sc_ind code
YOS_PG_count_XFXXX	Inventory count for all individuals in a given YOS, PG with sc_ind = XFXXX (active duty)
YOS_EMC_count_all	Inventory count for all individuals in a given YOS, EMC regardless of sc_ind code
YOS_EMC_count_XFXXX	Inventory count for all individuals in a given YOS, EMC with sc_ind = XFXXX (active duty)
PG_EMC_count_all	Inventory count for all individuals in a given PG, EMC regardless of sc_ind code
PG_EMC_count_XFXXX	Inventory count for all individuals in a given PG, EMC with sc_ind = XFXXX (active duty)

Data Element	Description
YOS_PG_EMC_count_all	Inventory count for all individuals in a given YOS, PG, EMC regardless of sc_ind code
YOS_PG_EMC_count_XFXXX	Inventory count for all individuals in a given YOS, PG, EMC with sc_ind = XFXXX (active duty)
sc_ind	XFXXX = active duty, reference EMF for definitions of other codes.
Comment	Comment denoting type of change observed in data

Appendix B

Observations in the data were deleted if YOS for a given PG exceeded the rules, provided below:

E2 (ACTIVE AND FTS) – 4 Years

E3 (ACTIVE AND FTS) - 5 years

E-4 - 10 years

*E-5 - 14 years (20 years for Reserves)

E-6 - 20 years

E-7 - 24 years

E-8 - 26 years

E-9 - 30 years

*The Navy changed E-5 HYT from 20 years to 14 years, effective July 1, 2005. However, Sailors with more than ten years of service as of July 1, 2005 may remain in the service until they are retirement eligible (20 years of service). For details, see NAVADMIN 056/05.

Appendix C

AME F-Test Two-Sample for Variances

	AR Forecast	MA-3 Forecast
Mean	0.687243953	0.69012292
Variance	0.005049067	0.00265388
Observations	338	350
df	337	349
F	1.902522747	
P(F<=f) one-tail	1.68921E-09	
F Critical one-tail	1.194515769	

AME t-Test: Two-Sample Assuming Unequal Variances

	MA-3 Forecast	AR Forecast
Mean	0.69012292	0.687243953
Variance	0.00265388	0.005049067
Observations	350	338
Hypothesized Mean Difference	0	
df	613	
t Stat	0.606662161	
P(T<=t) one-tail	0.272149873	
t Critical one-tail	1.647343168	
P(T<=t) two-tail	0.544299746	
t Critical two-tail	1.963841444	

t-Test: Two-Sample Assuming Equal Variances

	AME 3-Year Combination	AME Pseudo-Bayesian
Mean	0.036360598	-0.006290074
Variance	0.019158793	0.027768222
Observations	49	59
Pooled Variance	0.023869613	
Hypothesized Mean Difference	0	
df	106	
t Stat	1.42828657	
P(T<=t) one-tail	0.078074639	
t Critical one-tail	1.659356034	
P(T<=t) two-tail	0.156149277	
t Critical two-tail	1.982597262	

Appendix D

PG	YOS	Lag Order
E1-E3	0	0
E1-E3	1	0
E1-E3	2	4
E1-E3	3	0
E1-E3	4	2
E1-E3	5	1
E4	1	0
E4	2	0
E4	3	6
E4	4	0
E4	5	0
E4	6	7
E4	7	3
E5	3	3
E5	4	0
E5	5	2
E5	6	8
E5	7	4
E5	8	4
E5	9	5
E5	10	0
E5	11	0
E5	12	0
E5	13	0
E5	15	3
E5	16	1
E6	6	0
E6	9	2
E6	10	2
E6	11	7
E6	12	4
E6	13	5
E6	14	4
E6	15	4
E6	16	0
E6	17	4
E6	18	6
E6	19	0
E7	14	3
E7	15	0
E7	16	4
E7	17	1
E7	18	1
E7	19	1
E7	20	2

Distribution List

AIR UNIVERSITY LIBRARY
ARMY RESEARCH INSTITUTE LIBRARY
ARMY WAR COLLEGE LIBRARY
CENTER FOR NAVAL ANALYSES LIBRARY
HUMAN RESOURCES DIRECTORATE TECHNICAL LIBRARY
JOINT FORCES STAFF COLLEGE LIBRARY
MARINE CORPS UNIVERSITY LIBRARIES
NATIONAL DEFENSE UNIVERSITY LIBRARY
NAVAL HEALTH RESEARCH CENTER WILKINS BIOMEDICAL LIBRARY
NAVAL POSTGRADUATE SCHOOL DUDLEY KNOX LIBRARY
NAVAL RESEARCH LABORATORY RUTH HOOKER RESEARCH LIBRARY
NAVAL WAR COLLEGE LIBRARY
NPRST SPISHOCK LIBRARY
OFFICE OF NAVAL RESEARCH (CODE 34)
PENTAGON LIBRARY
US AIR FORCE ACADEMY LIBRARY
US COAST GUARD ACADEMY LIBRARY
US MERCHANT MARINE ACADEMY BLAND LIBRARY
US MILITARY ACADEMY AT WEST POINT LIBRARY
US NAVAL ACADEMY NIMITZ LIBRARY